

What is claimed is:

1. A method for alignment of a plurality of data traces indicative of the positions of a plurality of nucleic acid base types in a target nucleic acid, comprising the steps of:

(a) selecting a set of three or more alignment points for each data trace, said alignment points being selected from the group consisting of a primer peak associated with unextended primer, a full-length peak associated with full length product produced during a cyclic primer extension reaction with two primers, and an internal peak associated with internal bases that are highly conserved in the target nucleic acid, and said alignment points each having a reference position number reflecting the position of the alignment point with respect to the sequence as a whole;

(b) assigning a sequence position number to each peak in each of the plurality of data traces, said sequence position numbers being assigned to maximize the number of times that the sequence position number and the matching reference position number are assigned to a base of the same type; and

(c) aligning the data traces based on the assigned sequence position numbers.

2. The method of claim 1, wherein at least some of the internal peak alignment points are members of heterogeneous multiplets.

3. The method of claim 2, wherein all of the internal peak alignment points are members of heterogeneous multiplets.

4. The method of claim 1, wherein four data traces, one for each nucleotide base type, are aligned.

5. A method for alignment of a plurality of data traces indicative of the positions of a plurality of nucleic acid base types in a target nucleic acid, comprising the steps of:

- (a) selecting a set of five or more alignment points for each data trace, said alignment points being selected from the group consisting of a primer peak associated with unextended primer, a full-length peak associated with full length product produced during a cyclic primer extension reaction with two primers, and an internal peak associated with internal bases that are highly conserved in the target nucleic acid, and said alignment points each having a reference position number reflecting the position of the alignment point with respect to the sequence as a whole;
- (b) assigning a sequence position number to each peak in each of the plurality of data traces, said sequence position numbers being assigned to maximize the number of times that the sequence position number and the matching reference position number are assigned to a base of the same type; and
- (c) aligning the data traces based on the assigned sequence position numbers.

6. The method of claim 5, wherein at least some of the internal peak alignment points are members of heterogeneous multiplets.

7. The method of claim 5, wherein all of the internal peak alignment points are members of heterogeneous multiplets.

8. The method of claim 5, wherein four data traces, one for each nucleotide base type, are aligned.

9. An apparatus for determining the positions of nucleotide bases in a target nucleic acid, comprising:

- (a) a DNA sequencer comprising an electrophoresis system and a detection system for acquiring data traces reflecting the positions of nucleic acid bases in the target nucleic acid;

VGEN.P-066
PATENT APPLICATION

- (b) a computer connected to the sequencer to receive the data traces, said computer comprising a processor; and
 - (c) a storage device operatively connected to the computer in which inner alignment points and associated reference position numbers for one or more target nucleic acids are stored, wherein the processor is programmed to receive the data traces, access the inner alignment points and associated reference position numbers, identify peaks in the data traces and assign sequence positions numbers to the identified peaks of the data traces which maximize the number of times that the sequencing position number and the matching reference position number are assigned to a base of the same type.